Math 6020-1, Spring 2014; Assignment 3

Assigned on Friday February 7, 2014; Due on Friday February 27, 2014

- 1. Finish reading Chapter 8.
- 2. Solve (p. 475) problem 8.16
- 3. Download the "Passenger Car Mileage" data file from Carnegie Mellon's DSL library (html link). This is based on 1991 EPA emissions data, and is publicly available.
 - (a) Use the data to write a linear model, for predicting the fuel efficiency of a car, as follows:

$$MPG = \beta_0 + \beta_1 VOL + \beta_2 HP + \beta_3 SP + \beta_4 WT + \epsilon;$$

where MPG = miles per gallon; VOL = passenger cab volume in cubic feet; HP = engine horse power; SP = top speed in mph; WT = the vehicle weight in hundreds of pounds; and ϵ = noise.

- (b) Find LSE estimates for $\boldsymbol{\beta} := (\beta_0, \dots, \beta_4)'$. Apply confidence intervals and/or ellipsoids in order to decide which, if any, of the explanatory variables are redundant. What assumptions are you making about the noise?
- (c) What is your least-squares predictor of the MPG of a car whose vital statistics are: VOL= 91 ft³; HP= 85; SP= 72 mph; and WT= 5320 pounds?
- (d) Let X denote the vector (VOL, HP, SP, WT) that corresponds to a randomly-selected vehicle in 1991. How would you estimate $\Sigma :=$ Cov(X)? Perform this estimation, and explain [in as much detail as you can] your reasoning.
- (e) Estimate the correlation matrix ρ . Describe your estimator carefully, and explain why you think it is a reasonable estimator.
- (f) Carry out a PCA on Σ and a PCA on ρ .
- (g) Can you identify what the first one or two principle components might designate in each case?
- (h) Which PCA do you recommend here? Explain your reasoning carefully.